

WE CLAIM:

1 1. In a system including spatial data for a spatial environment, wherein a recipe
2 is used in the spatial environment, a method for mining the spatial data to optimize the
3 recipe for one or more target values, the method comprising:

4 an act of generating a data set from the spatial data using identified attributes
5 selected by a user;

6 an act of inspecting the generated data set to provide statistical information
7 for the data set;

8 an act of preprocessing the data set to prepare the data set for modeling;

9 an act of modeling the preprocessed data set to describe relationships
10 between the attributes and the one or more target values; and

11 an act of providing recommendations such that the recipe is optimized.

1 2. A method as defined in claim 1, wherein the act of preprocessing the data
2 set further comprises:

3 an act of cleaning the generated data set;

4 an act of interpolating the generated data set;

5 an act of normalizing the generated data set; and

6 an act of generating new attributes.

1 3. A method as defined in claim 1, wherein the recipe is a fertilizer recipe
2 for use in an agricultural field.

3 4. A method as defined in claim 1, wherein a crop yield is included in the
4 one or more target values.

1 5. A method as defined in claim 1, further wherein the relationships include
2 one or more clusters, wherein a first cluster from first spatial data corresponding to a
3 first spatial environment is used to optimize a recipe for a second spatial environment.

1 6. A computer program product having computer executable instructions for
2 executing the acts recited in claim 1.

1 7. In a system including one or more spatial databases corresponding to one
2 or more spatial environments, a system for knowledge discovery from the one or more
3 spatial databases, the system comprising:

4 a user interface; and

5 a spatial data modeling and analysis module (SDAM module) for
6 extracting knowledge from the one or more spatial databases, the SDAM module
7 comprising:

8 a data generation and manipulation module for loading the data
9 set from the one or more spatial databases based on designated attributes,
10 wherein attributes are supplied to the data generation and manipulation
11 module by a user through the user interface;

12 a data inspection module for providing spatial statistics on the
13 loaded data set;

14 a data preprocessing module for preparing the data set for
15 modeling, wherein the data preprocessing module removes errors from
16 the data set;

17 a data partitioning module for dividing the data set into
18 homogenous data segments which improve data modeling; and

19 a modeling module for describing relationships between the
20 attributes and one or more target values, wherein the relationships are
21 obtained from the partitioned data set.

1 8. A system as defined in claim 7, wherein the SDAM module further
2 comprises an integration module for enhancing the knowledge generated from the one or
3 more spatial databases.

1 9. A system as defined in claim 7, wherein the preprocessing module further
2 comprises:

3 a cleaning and filtering module for removing duplicate data and removing
4 noise from the loaded data set;

5 a data interpolation module for computing common values for a common
6 set of locations;

12 a data inspection module for providing spatial statistics on the
13 loaded data set;

14 a data preprocessing module for preparing the data set for
15 modeling, wherein the data preprocessing module removes errors from
16 the data set;

17 a data partitioning module for dividing the data set into
18 homogenous data segments which improve data modeling; and

19 a modeling module for describing relationships between the
20 attributes and one or more target values, wherein the relationships are
21 obtained from the partitioned data set.

1 8. A system as defined in claim 7, wherein the SDAM module further
2 comprises an integration module for enhancing the knowledge generated from the one or
3 more spatial databases.

1 9. A system as defined in claim 7, wherein the preprocessing module further
2 comprises:

3 a cleaning and filtering module for removing duplicate data and removing
4 noise from the loaded data set;

5 a data interpolation module for computing common values for a common
6 set of locations;

7 a data normalization module for transforming the loaded data set to a
8 normal distribution and for scaling the loaded data set to a range;
9 a data discretization module for use in modeling the loaded data set;
10 a generating new attributes module for combining existing attributes into
11 a single attribute;
12 a feature selection module for reducing the attributes identified by a user
13 such that irrelevant attributes may be removed; and
14 a feature extraction module for reducing a dimensionality of the loaded
15 data set.

1 10. A system as defined in claim 7, further comprising a recommendation
2 module, wherein the recommendation module optimizes a recipe for a spatial
3 environment.

1 11. A system as defined in claim 10, wherein the recommendation module
2 includes at least one of: a fertilization module for optimizing a fertilizer recipe to be
3 applied to an agricultural field; an irrigation module for optimizing a water recipe to be
4 applied to a field; and an equipment module for optimizing a recipe to be applied to
5 equipment.

1 12. A system as defined in claim 11, wherein the recommendation module
2 includes at least one of: a pesticide module, a herbicide module, and a seed-spacing
3 module.

1 13. A system as defined in claim 7, wherein each of the data generation and
2 manipulation module, the data inspection, the data preprocessing module, the data
3 partitioning module, and the modeling module can be independently controlled by the
4 user through the user interface.

1 14. In a networked computer system that includes a client and a server, wherein
2 the server maintains spatial data sets, a method for analyzing the spatial data sets over the
3 network, the method comprising the steps for:

4 applying spatial data mining functions to the spatial data sets, wherein said
5 spatial data mining functions comprise the steps for

6 modeling the spatial data sets to provide estimation of predetermined
7 parameters at predetermined points; and

8 classifying the spatial data sets into predetermined classes; and

9 using the estimation of the predetermined parameter to accomplish a
10 predetermined purpose, wherein the predetermined purpose includes at least one of
11 determining how the predicted variable affects a predetermined target variable,
12 providing recommendations as to how to achieve a predetermined target variable,
13 and creating new spatial data mining methods.

1 15. A method as defined in claim 14, further comprising the step for combining
2 different programming environments to allow different programming environments to
3 function on one server.

1 16. A method as defined in claim 15, wherein the step for combining different
2 programming environments comprises a unified controller.

1 17. A method as defined in claim 14, wherein the spatial data set is generated
2 by a spatial data simulator.

1 18. A method as defined in claim 14, wherein said spatial data mining functions
2 further comprise the step for partitioning said data set into more homogenous portions.

1 19. A method as defined in claim 14, wherein said spatial data mining functions
2 further comprise the step for integrating said modeling and classifications steps.

1 20. A computer program product having computer executable instructions for
2 performing the steps recited in claim 14.

1 21. In an environment including spatial data relating to a specific agricultural
2 field, a method for analyzing the spatial data comprising steps for:

3 applying spatial data mining functions to the spatial data, wherein said
4 spatial data mining functions comprise the steps for
5 modeling the spatial data to provide estimation of predetermined parameters
6 at predetermined points; and
7 classifying the spatial data into predetermined classes;
8 using the results of the spatial data analysis to optimize the treatment
9 of the agricultural field to produce a predetermined yield.

1 22. A method as defined in claim 21, wherein said spatial data consists of past
2 and present data of a specific agricultural field.

1 23. A method as defined in claim 21, wherein the step for applying spatial data
2 mining functions occurs in a network environment.